

# Compensating lung cancer patients occupationally exposed to coal tar pitch volatiles

Ben Armstrong, Gilles Thériault

## Abstract

**Objectives**—To investigate the process of deciding on compensation claims by lung cancer patients exposed occupationally to coal tar pitch volatiles.

**Methods**—For each case of lung cancer the probability that it was caused (probability of causation (PC)) by coal tar pitch volatiles was expressed as an increasing function of cumulative exposure to benzo(a)pyrene-years. This was assessed from several exposure-response models fitted to data from a large epidemiological study of aluminium production workers. For some models, PC depended also on the smoking habit of the cancer patient.

**Results**—Estimation of relative risk by exposure group indicated that over 50% of lung cancers were attributable to coal tar pitch volatiles (PC >50%) at exposures above 100  $\mu\text{g}/\text{m}^3$ -years benzo(a)pyrene. A linear relative risk model indicated that 50% PC was first achieved at 342.2  $\mu\text{g}/\text{m}^3$ -years benzo(a)pyrene, or 190.1  $\mu\text{g}/\text{m}^3$ -years benzo(a)pyrene according to the upper 95% confidence limit for risk increment. Corresponding figures for a power curve model were 210.3 and 45.9. With these five figures as compensation criteria compensation would have resulted in 31.4%, 2.7%, 19.2%, 15.7%, and 39.2% of cancers studied, compared with an estimated total proportion of cancers studied attributable to coal tar pitch volatiles of 15%–26%. If risks due to coal tar pitch volatiles and smoking multiply, PC does not depend on the amount smoked. If the two risks are additive, however, PC depends on the amount smoked according to a formula, with the figures mentioned applying to an average smoking history (24.4 pack-years).

**Conclusion**—Because of its simplicity and because it falls within the range of criteria based on several more sophisticated approaches, we prefer the criterion of 100  $\mu\text{g}/\text{m}^3$ -years, based on the relative risks by exposure group. However, the compensation board of the Canadian province of Quebec, on consideration of these alternatives, has proposed as a criterion that the upper 95% confidence limit of PC for the patient be at least 50%, assuming an additive relative risk model and allowing for their smoking habit.

**Keywords:** coal tar pitch volatiles; compensation; models

The production of aluminium metal from alumina by the Soderberg process has in the past entailed substantial exposures to coal tar pitch volatiles, and in particular polynuclear aromatic hydrocarbons (PAHs). Several epidemiological studies have associated these exposures in this industry with an excess risk of lung cancer.<sup>1</sup> We have recently reported a further epidemiological study that showed an association of coal tar pitch volatiles with lung cancer in Soderberg workers from Quebec,<sup>2</sup> not due to confounding by smoking. The study also estimated the exposure-response relation between coal tar pitch volatiles and lung cancer.

In this paper we present evidence from the epidemiological study that is relevant to the deciding criteria for whether to compensate lung cancer victims among aluminium production plant workers, and possibly others exposed to PAHs. This compliments our earlier discussion on compensation of bladder cancer victims in the same industry.<sup>3</sup>

## Methods

### USE OF EPIDEMIOLOGICAL DATA IN DECIDING COMPENSATION

Workers compensation schemes were set up primarily for the compensation of workers experiencing occupational injuries. For these cases, the cause of the injury is rarely a matter of dispute. This is also true for diseases specific to the occupation, such as pneumoconiosis or occupational asthma. Increasingly, diseases with occupational and non-occupational causes, and which have no objective features specific to the occupational cause, have been the object of compensation claims. Examples are lung cancer among workers exposed to asbestos or to radioactive gases, and bladder cancer among workers exposed to certain chemicals used in dye and rubber manufacturing. For these diseases, evidence for an occupational cause rests on evidence from epidemiological studies, and is in the form of statements on increased risks in groups of exposed workers. It is not possible to identify with certainty which specific workers have had cancer caused by work exposures.

Occasionally, it is possible to assess causes, with some assumptions, on the basis of the presence of other pathologies assumed to lie on the causal path between the occupational

McGill University  
Department of  
Occupational Health,  
1130 Pine Avenue  
West, Montreal,  
Quebec, Canada  
H3A 1A3

B Armstrong  
G Thériault

Current address:  
Environmental  
Epidemiology Unit,  
London School of  
Hygiene and Tropical  
Medicine, Keppel  
Street, London  
WC1E 7HT

B Armstrong  
Correspondence to:  
Dr B Armstrong,  
Environmental Epidemiology  
Unit, London School of  
Hygiene and Tropical  
Medicine, Keppel Street,  
London WC1E 7HT.

Accepted 6 October 1995

(*Occup Environ Med* 1996;53:160–167)

exposure and the cancer. For example, some compensation authorities compensate people with lung cancers who had occupational exposure to asbestos only on evidence of asbestosis.

Compensation is not usually considered unless a causal association between work and a cancer has been established beyond reasonable doubt for groups. The criteria by which this is done is the subject of epidemiological text books and a substantial number of journal articles (Rothman<sup>4</sup>). This is not the subject of this review. We assume that a causal association is established for groups of exposed workers, and discuss whether and how epidemiological data can help in deciding whether to compensate specific cases.

#### PROBABILITY OF CAUSATION

The concept most often used in this context is that of probability of causation (PC).<sup>5</sup> The PC can be thought of as the chance that a specific case of cancer was caused by the exposure in question. More formally, it is defined from the risks of cancer in people exposed ( $R_e$ ) and unexposed ( $R_u$ ) (both exposed and unexposed having other risk factors the same as the case) as:

$$PC = (R_e - R_u)/R_e$$

the excess risk divided by the total risk in exposed people. By dividing top and bottom of this fraction by  $R_u$ , we can express the PC in terms of the relative risk (RR):

$$RR = R_e/R_u \text{ (of exposed } v \text{ unexposed):}$$

$$PC = ((R_e/R_u) - (R_u/R_u))/(R_e/R_u) = (RR - 1)/RR \quad (1)$$

Note that this means that at a relative risk of two, the PC is 50%, as is intuitively reasonable. The PC is known to epidemiologists as the aetiological fraction, causal fraction, attributable risk for exposed persons, or attributable proportion and is equal to the proportion of excess cases in the exposed population relative to the unexposed population.

The RR in formula (1), and hence the PC, will usually depend on the quantity of exposure. Sometimes, the RR and PC depend also on other factors, sometimes they do not. For example, relative risk of lung cancer after a given exposure to asbestos is usually supposed to be independent of age, even though absolute risk increases with age. We say that the risks due to age and to asbestos multiply. If this is so, we do not need to know cases' ages to be able to calculate their PCs. However, sometimes risks do not multiply, so that the RR is different for different ages. When this is true, we need to know the cases' ages to calculate their PCs.

This argument is true also when the second risk factor is avoidable, such as smoking. For example, assume that the lifetime risk of lung cancer in a non-smoker is 1%, and that of a smoker is 10%. Assume further that an occupational exposure doubles the risk ( $RR = 2$ ) for both smokers and non-smokers (the risks multiply). Then for non-smokers,  $R_u = 1\%$ ,  $R_e = 2\%$ , and for a non-smoking case the PC from occupational exposure is  $(2 - 1)/2 =$

50%. For smokers,  $R_u = 10\%$ ,  $R_e = 20\%$ , and for a smoking case the PC from occupational exposure is  $(20 - 10)/20 = 50\%$ , the same as for a non-smoking case  $((2 - 1)/2)$ . For a smoking case, the PC for smoking is  $(20 - 2)/20 = 90\%$ . The PCs for occupational exposure and smoking thus added are  $50 + 90 = 140\%$ . This result, which is sometimes found counter intuitive, is nevertheless well established.<sup>6</sup> In 100 workers, there would be 20 cancers if they smoked and were exposed to the occupational risk factor, 10 if they smoked but were not exposed to the occupational risk factor, 2 if they were exposed to the occupational risk factor but did not smoke, and 1 if they were exposed to neither. Given multiplication of risks, the combined action of both contaminants is more than the sum of each.

However, where the risks do not multiply, the RR and PC will depend on the value of the second risk factor. For example, an occupational exposure might add to rather than multiply with background age specific risks of lung cancer. Then, because lung cancer rates rise with age, the PC will be higher in a young case than an old case, if the exposures are the same in the two. Similarly, if risks due to smoking and those due to the occupational exposure are additive, the PC will depend on whether and how much a case smoked. For example, assume that the background lifetime risk of dying from lung cancer in non-smokers was 1% as before, the occupational exposure added 10% and smoking added 9%. Then for smokers,  $R_u = 10\%$ ,  $R_e = 20\%$ ,  $RR = 20/10 = 2$ , and for a smoking case the  $PC = (20 - 10)/20 = 50\%$ , as in the previous example, in which the risks multiplied. However, for non-smokers,  $R_u = 1\%$ ,  $R_e = 11\%$ ,  $RR = 11$ , and for a non-smoking case the  $PC = (11 - 1)/11 = 91\%$ , much higher than in the previous example, and higher than for a smoking case.

#### USE OF THE PC IN DECIDING COMPENSATION

The definition and properties of PCs are the concern of science. Whether and how PCs are used in deciding compensation is as much a matter for the legal, social, and political process as it is for science.

The "preponderance of evidence" rule in law courts in the United States for plaintiffs in civil "toxic tort" suits formally requires demonstration that the PC for a cancer is greater than 50%.<sup>7</sup> This is contrasted with the "beyond reasonable doubt" criterion required for a conviction in a criminal case. In practice, individual plaintiffs have apparently often been successful even where the PC has been less than 10%, apparently due to the judges evaluating the probability that the exposure had contributed to rather than caused cancer.<sup>8</sup>

The same principle is often used, usually crudely, in deciding whether to compensate workers. Thus, where a causal association with  $RR > 2$  has been established, workers are considered eligible for compensation. This is often on the basis of simple tables of RR by duration of exposure in an exposed depart-

ment. However, where there are exposure summaries such as cumulative exposure, which explain risk better, they are often used.<sup>9</sup>

The use of RR estimates in subjects grouped by exposure have the advantage of simplicity, but may often depend critically on the choice of group boundaries, which are arbitrary. The use of mathematical models for RR avoids this, at the expense of loss of simplicity.

Initially an automatic compensation criterion was set for cases of bladder cancer among aluminium production workers in Quebec at that exposure that was estimated by a model to give a PC of 50%. Here, to give the "benefit of the doubt" to the worker, the 95% upper confidence limit (UCL) for the risk estimate was used to calculate this value.<sup>3</sup> The workers compensation board of British Columbia also chose to use a PC of 50% as a compensation criterion. However, they opted to use that based on the point estimate for risk, rather than the confidence limit.

For gold miners in Ontario, the Ontario Industrial Disease Standards Panel proposed ranking cases according to a cumulative exposure index and hence PC. However, rather than choosing the minimum exposure that would result in compensation as that giving 50% PC, they proposed a level that made the number of cases above that level in the epidemiological study equal to the excess estimated (using the 95% UCL) to be due to exposure. In fact this coincided with an RR of about 1.4, and hence a PC of  $0.4/1.4 = 29\%$ .<sup>10</sup>

For uranium miners in Ontario, the Industrial Diseases Standards Panel proposed as a criterion, cumulative exposure to radioactive gases that gave an RR of 2 (PC of 50%). They allowed benefit of the doubt to the worker in favourable assumptions of past exposures, where these were in doubt.<sup>9</sup>

The best known direct use of the PC has been in compensation decisions on cancers related to radiation. The United States Congress 1983 Orphan Drug Act required the National Institutes of Health to develop tables of the PCs for use in compensation decisions.<sup>11</sup> These tables showed PC as a function of total dose of radiation. Because the association of risk and dose varies by anatomical site, separate tables were produced for each cancer site. Under this law, if the PC is less than 10%, claimants are not compensated, and if the PC is greater than 50%, claimants receive full compensation. For a PC between 10% and 50%, claimants receive compensation equal to their PC times the full compensation sum. For example, someone with a PC = 30% would receive 30% of the amount received by someone with a PC >50%. Such proportional compensation schemes have attractive features for use in workers' compensation, in particular avoiding a cut off point above which a claimant gets a large sum of money, and below which they get nothing.<sup>12</sup> However, we are not aware of any application by workers' compensation boards.

#### CRITIQUES OF THE PC AS A COMPENSATION CRITERION

Chase *et al*<sup>13</sup> discuss eligibility for compensation in lung cancer patients among workers exposed to asbestos. They object to an approach based on the PC because, under the multiplicative model for the combined effect of asbestos and smoking (which they assumed), the PC approach treats smokers and non-smokers alike. They argue that because much of the additional absolute risk due to asbestos in smokers is due to the synergism of asbestos with smoking, smokers have, by choosing to smoke, augmented the risk per unit asbestos exposure. They propose that compensation be based on a concept that they call "risk apportioned to asbestos" (RAA). RAA is defined as:

$$RAA = ((a - 1)(s - 1))/(as(a + s - 2)) \quad (2)$$

where  $a$  is the RR due to asbestos alone, and  $s$  is the RR due to smoking alone. For a non-smoker,  $s = 1$ , and  $RAA = (a - 1)/a = PC$ . For a smoker,  $s > 1$ , and hence  $RAA < PC$ , by an amount that depends on the level of smoking. They suggest that in calculating  $s$ , smoking before the United States Surgeon General's first clear warning that smoking was harmful to health (taken as 1966) should not be counted.

We allow the definition (1) of the PC to apply when the RR refers to ratios of incidence rates (hazard ratios) as well as ratios of risks in the strict sense of probabilities of contracting the disease or dying from it in a fixed period of time. Robins and Greenland<sup>14-17</sup> argue that this extension does not allow the interpretation that the PC represents the proportion of cases that would not have occurred over a period of time without the exposure. We nevertheless feel confident in using PCs with this interpretation here because in this population the hazard ratios will be very similar to ratios of lifelong risk of lung cancer. This is so because lung cancer remains a rare disease. We need not enter into the debate as to usefulness of PCs when hazard ratios do not approximate to lifelong RRs.

#### RELATION BETWEEN COAL TAR PITCH VOLATILES AND LUNG CANCER IN ALUMINIUM PRODUCTION WORKERS

We concluded from our epidemiological study<sup>2</sup> and the sum of earlier knowledge, that there had clearly been an excess of lung cancer in the plant studied that was caused by occupational exposures, and that coal tar pitch volatiles were easily the most likely specific cause. The arguments leading to this conclusion have been published previously,<sup>2</sup> and are not repeated here.

The study showed clearly increased risks, after controlling for smoking, in Soderberg potroom workers, and suggested increased risks in workers in other departments in which there was some exposure to coal tar pitch volatiles. There was no evidence for a threshold level of coal tar pitch volatiles below which the risk was zero, and no such threshold would generally be expected for a carcinogen. The

data were compatible with a linear relation of RR with coal tar pitch volatiles, but showed some tendency towards a steeper slope at lower exposures, which was represented mathematically as a power curve model. Coal tar pitch volatiles were measured as benzene soluble matter and as benzo(a)pyrene. The association with lung cancer was similar for benzene soluble matter and for benzo(a)pyrene, but as benzo(a)pyrene was initially expected to be more closely related to the causal agents, PAHs, we concentrate our attention on this.

There was no evidence in the data for a minimum period of time (latency or induction time) before which exposure affected risk. Nevertheless, general knowledge on mechanisms of the action of chemicals on occurrence of cancer suggests that the exposure in the immediate five years before death was unlikely to have affected risk. In the mathematical models for risk, we do not count this exposure.

The best fitting linear model for RR ( $RR(x)$ ) of a subject exposed to  $x \mu\text{g}/\text{m}^3$ -years of coal tar pitch volatiles is given by:

$$RR(x) = 1 + \beta x \quad (3)$$

with  $\beta = 0.002922$  (95% confidence interval (95% CI)  $0.001168$ – $0.005261$ )<sup>2</sup>. According to the alternative power curve model:

$$RR(x) = 1 + \beta x^p \quad (4)$$

with  $p = 0.380$ , and  $\beta = 0.131$  (95% CI  $0.060$ – $0.234$ ,  $p$  fixed). For example, at  $200 \mu\text{g}/\text{m}^3$ -years of benzo(a)pyrene, the RR is estimated at  $1 + (0.0029 \times 200) = 1.58$ , (95% CI  $1.24$ – $2.06$ ) with the linear model, and  $1 + (0.131 \times 200^{0.380}) = 1.98$ , (95% CI  $1.45$ – $2.75$ ) with the power curve model. Both models were compatible with the data.

From analyses of data on the smoking habits of workers, we found that non-exposed workers smoked to the same extent as exposed workers. Thus smoking did not confound the relation between coal tar pitch volatiles and lung cancer. Also, the data were consistent with the hypothesis that the relation (2) applied equally to smokers and non-smokers (although the baseline risk is obviously less for non-smokers). This is the multiplicative model for the combined effect of smoking and coal tar pitch volatiles. However, the data were also consistent with the additive model, under which the RR is higher for non-smokers than smokers. The appendix gives mathematical details of these two models, with implications for calculating PC.

The study was of mortality from lung cancer, whereas presumably incident cases will apply for compensation. Because of the high mortality of lung cancer, statements that we make about RR and PC for lung cancer deaths will also apply as a very good approximation to incident cases.

#### ESTIMATING PCS FOR THE ALUMINIUM PRODUCTION WORKERS

To estimate a PC from tables of RRs in a group—for example,  $10$ – $19 \text{ mg}/\text{m}^3$ -years benzene soluble matter—we simply calculate  $(1 -$

$RR)/RR$ . From the mathematical models, the PC is given by expression (1):

$$PC(x) = (RR(x) - 1)/RR(x)$$

As we also assume that  $RR(x) = 1 + \beta x$ , we can re-express this as

$$PC(x) = \beta x / (1 + \beta x) \quad (5)$$

An estimate and 95% CI for  $\beta$  are already given.

This formula ignores smoking. As we have noted above, if the risks of smoking and exposure to coal tar pitch volatiles multiply, the same formula applies to all subjects, whatever their smoking history. However, data from the study, and existing knowledge of the action of carcinogens, are compatible with either a multiplicative or additive combination of risks due to smoking and coal tar pitch volatiles. We therefore calculate two tables of PCs, one assuming the multiplicative model equation (5), one assuming the additive model, where (see appendix):

$$PC(x) = 9.05\beta x / (1 + (0.33 \times \text{pack-years}) + 9.05\beta x) \quad (6)$$

#### Results

Before proceeding to estimate the PC, we use the published results from the study to estimate the total fraction of cases in the study attributable to exposure to coal tar pitch volatiles. This global attributable fraction may be thought of as a PC for a case for which the only information was that he was in the study. However, we do not use this information directly to decide a compensation criterion, but as a benchmark with which to compare criteria that use other estimates of the PC which differentiate between workers.

A simple estimate of this global attributable fraction is obtained by estimating risk in all exposed workers ( $>10 \mu\text{g}/\text{m}^3$ -years benzo(a)pyrene), relative to the unexposed workers. This RR is  $1.75$  ( $1.41$ – $2.19$ ), implying that the proportion of exposed cases attributable to their exposure was  $(1 - 1.75)/1.75 = 42.9\%$  (95% CI  $29.1$ – $54.3$ ). Because  $59.2\%$  of cases were exposed, this gives an attributable fraction in all workers combined of  $59.2 \times 42.9 = 25.4\%$  (95% CI  $17.7$ – $32.2$ ). Summing the attributable numbers from the table of RR by grouped exposure<sup>2</sup> gives the same estimate. The estimated attributable fraction with the linear RR model is  $14.7\%$  ( $7.8$ – $20.9$ ), and with the power curve model is  $20.4\%$  ( $16.6$ – $38.4$ ).

One way to set a compensation criterion is to choose a minimum exposure that would allow compensation of the right proportion of cases in the study. For example, to have compensated  $25.4\%$  of cases would require a criterion of  $141 \mu\text{g}/\text{m}^3$ -years benzo(a)pyrene.

The simplest approach to estimating the PC in a manner that differentiates between workers is to use the analysis of RR according to duration of work in each department.<sup>2</sup> This shows, after adjustment for confounding, RRs in excess of two (PC  $>50\%$ ) for workers work-

Table 1 Probability (%(95% CIs)) that lung cancer was caused by exposure to coal tar pitch volatiles (benzo(a)pyrene): linear model

Benzo(a) pyrene ( $\mu\text{g}/\text{m}^3\text{-y}$ )	Multi- plicative model	Additive model: by smoking (pack-years)								
		0	10	20	30	40	50	60	70	80
0	0-0	0-0	0-0	0-0	0-0	0-0	0-0	0-0	0-0	0-0
	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)
50	12-7	56-9	23-5	14-8	10-8	8-5	7-0	6-0	5-2	4-6
	(5-5, 20-8)	(34-6, 70-4)	(10-9, 35-6)	(6-5, 23-9)	(4-6, 17-9)	(3-6, 14-4)	(2-9, 12-0)	(2-5, 10-3)	(2-1, 9-0)	(1-9, 8-0)
100	22-6	72-6	38-1	25-8	19-5	15-7	13-1	11-3	9-9	8-8
	(10-5, 34-5)	(51-4, 82-6)	(19-7, 52-5)	(12-2, 38-5)	(8-8, 30-4)	(6-9, 25-1)	(5-7, 21-4)	(4-8, 18-6)	(4-2, 16-5)	(3-7, 14-8)
150	30-5	79-9	48-0	34-3	26-7	21-8	18-5	16-0	14-1	12-6
	(14-9, 44-1)	(61-3, 87-7)	(26-9, 62-4)	(17-3, 48-5)	(12-7, 39-6)	(10-0, 35-5)	(8-3, 29-0)	(7-1, 25-6)	(6-2, 22-9)	(5-5, 20-7)
200	36-9	84-1	55-2	41-0	32-7	27-1	23-2	20-3	18-0	16-2
	(18-9, 51-3)	(67-9, 90-5)	(33-0, 68-9)	(21-8, 55-6)	(16-2, 46-6)	(13-0, 40-1)	(10-8, 35-2)	(9-2, 31-4)	(8-1, 28-3)	(7-2, 25-8)
250	42-2	86-9	60-6	46-5	37-8	31-8	27-4	24-1	21-5	19-4
	(22-6, 56-8)	(72-6, 92-3)	(38-1, 73-5)	(25-8, 61-0)	(19-5, 52-62)	(15-7, 45-6)	(13-1, 40-5)	(11-3, 36-4)	(9-9, 33-1)	(8-8, 30-3)
300	46-7	88-8	64-9	51-1	42-1	35-8	31-2	27-6	24-8	22-5
	(25-9, 61-2)	(76-0, 93-5)	(42-5, 76-9)	(29-4, 65-3)	(22-5, 56-7)	(18-3, 50-2)	(15-3, 44-9)	(13-2, 40-7)	(11-6, 37-2)	(10-4, 34-3)
350	50-6	90-3	68-3	54-9	45-9	39-5	34-6	30-8	27-8	25-3
	(29-0, 64-8)	(78-7, 94-3)	(46-3, 79-5)	(32-7, 68-7)	(25-3, 60-5)	(20-7, 54-0)	(17-5, 48-8)	(15-1, 44-5)	(13-3, 40-9)	(11-9, 37-8)
400	53-9	91-4	71-1	58-2	49-3	42-7	37-7	33-7	30-5	27-9
	(31-8, 67-8)	(80-9, 95-0)	(49-6, 81-6)	(35-8, 71-5)	(28-0, 63-6)	(22-9, 57-3)	(19-5, 52-1)	(16-9, 47-8)	(14-9, 44-1)	(13-4, 41-0)

The exact parameter values used to obtain these figures are given in the appendix.

ing more than 20 years in the Soderberg potrooms (74 cases) or in the carbon plant (6 cases).

The next step in sophistication is to use the simple grouped analysis by cumulative exposure. This showed RR in excess of two (PC >50%) at 100  $\mu\text{g}/\text{m}^3\text{-years}$  benzo(a)pyrene.<sup>2</sup> Although the PC dips slightly below 50% in some of the higher groups, this suggests a possible minimum exposure criterion of 100  $\mu\text{g}/\text{m}^3\text{-years}$  benzo(a)pyrene (106 cases).

Finally, we use mathematical models for risk. Table 1 shows the estimated PC according to the linear models for exposure to coal tar pitch volatiles and smoking, given cumulative exposure to coal tar pitch volatiles and to cigarettes measured as pack-years (average number of packs a day  $\times$  years smoked). If we assume the multiplicative model (first column) the PC is the same regardless of how much the case smoked (see above). The estimate rises from zero for the unexposed workers to 53.9% at 400  $\mu\text{g}/\text{m}^3\text{-years}$  benzo(a)pyrene, reaching 50% between 300 and 350  $\mu\text{g}/\text{m}^3$  years. The CIs for PCs are also given. These show, for example that, at the upper limit for  $\beta$ , the PC reaches 50% at just below 200  $\mu\text{g}/\text{m}^3\text{-years}$  benzo(a)pyrene.

If we assume the additive model (remaining columns), the PC decreases, for a given expo-

sure to coal tar pitch volatiles, with the case's pack-years of smoking. Thus for a non-smoker, an exposure of only 50  $\mu\text{g}/\text{m}^3\text{-years}$  benzo(a)pyrene is estimated as quite likely to have been the cause of his cancer (PC = 56.9%), whereas for a heavy smoker (80 pack-years), an exposure of even 400  $\mu\text{g}/\text{m}^3\text{-years}$  benzo(a)pyrene is estimated as unlikely (PC = 27.9%) to have been the cause of his cancer.

Table 2 is analogous to table 1, except that it is based on a curved rather than linear model. According to this model, more workers would meet minimum exposure criteria based on the 50% PC than would under the linear models.

In table 3 we present the exposures for which a PC is 50%, according to the mathematical models. Two columns are given for each model, one uses the point estimate of  $\beta$ , and one uses the upper 95% confidence limit (UCL) of  $\beta$ , giving a lower 95% confidence limit (LCL) of the exposure. For example, under the linear additive model, a heavy smoker has to have 786.4  $\mu\text{g}/\text{m}^3\text{-years}$  benzo(a)pyrene to show an even (50%) chance that his cancer was caused by exposure to coal tar pitch volatiles (436.8 according to the UCL for  $\beta$ ). A non-smoker, on the other hand, has only to have 37.84  $\mu\text{g}/\text{m}^3\text{-years}$  benzo(a)pyrene (or 21.0 with the UCL). The

Table 2 Probability (% (95% CIs)) that lung cancer was caused by exposure to coal tar pitch volatiles (benzo(a)pyrene): power curve model

Benzo(a) pyrene ( $\mu\text{g}/\text{m}^3\text{-y}$ )	Multi- plicative model	Additive model: by smoking (pack-years)								
		0	10	20	30	40	50	60	70	80
0	0-0	0-0	0-0	0-0	0-0	0-0	0-0	0-0	0-0	0-0
	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)	(0-0, 0-0)
50	36-7	84-0	55-0	40-8	32-5	27-0	23-1	20-1	17-9	16-1
	(20-8, 50-8)	(70-4, 90-3)	(35-6, 68-5)	(23-9, 55-2)	(17-9, 46-2)	(14-4, 39-7)	(12-0, 34-8)	(10-3, 31-0)	(9-0, 28-0)	(8-0, 25-4)
100	43-0	87-2	61-4	47-3	38-5	32-5	28-1	24-7	22-1	19-9
	(25-5, 57-3)	(75-6, 92-4)	(41-9, 73-9)	(29-0, 61-5)	(22-1, 52-7)	(17-9, 46-1)	(15-0, 41-0)	(13-0, 36-9)	(11-4, 33-5)	(10-2, 30-7)
150	46-8	88-8	64-9	51-2	42-2	35-9	31-3	27-7	24-8	22-5
	(28-5, 61-0)	(78-3, 93-4)	(45-7, 76-7)	(32-2, 65-1)	(24-9, 56-6)	(20-3, 50-0)	(17-1, 44-8)	(14-8, 40-5)	(13-0, 37-1)	(11-6, 34-1)
200	49-5	89-9	67-4	53-9	44-9	38-5	33-7	29-9	26-9	24-5
	(30-8, 63-6)	(80-1, 94-1)	(48-4, 78-6)	(34-6, 67-6)	(27-0, 59-2)	(22-1, 52-7)	(18-7, 47-5)	(16-2, 43-2)	(14-3, 39-6)	(12-8, 36-6)
250	51-6	90-6	69-2	56-0	47-0	40-5	35-6	31-7	28-6	26-1
	(32-6, 65-5)	(81-4, 94-5)	(50-5, 80-0)	(36-6, 69-4)	(28-7, 61-2)	(23-6, 54-8)	(20-0, 49-6)	(17-4, 45-3)	(15-4, 41-7)	(13-8, 38-6)
300	53-4	91-2	70-7	57-7	48-7	42-2	37-2	33-2	30-1	27-4
	(34-2, 67-1)	(82-5, 94-9)	(52-2, 81-1)	(38-2, 70-8)	(30-1, 62-9)	(24-9, 56-5)	(21-2, 51-3)	(18-4, 47-0)	(16-3, 43-4)	(14-6, 40-2)
350	54-8	91-7	71-9	59-1	50-2	43-6	38-6	34-6	31-3	28-6
	(35-5, 68-4)	(83-3, 95-1)	(53-7, 82-0)	(39-6, 72-0)	(31-4, 64-2)	(26-0, 57-9)	(22-2, 52-8)	(19-3, 48-5)	(17-1, 44-8)	(15-4, 41-7)
400	56-1	92-0	72-9	60-3	51-5	44-9	39-8	35-7	32-4	29-7
	(36-7, 69-5)	(84-0, 95-4)	(54-9, 82-7)	(40-8, 73-0)	(32-5, 65-4)	(27-0, 59-2)	(23-0, 54-0)	(20-1, 49-7)	(17-9, 46-1)	(16-1, 42-9)

CIs for the power curve model are calculated assuming that the power parameter is fixed. The exact parameter values used to obtain these figures are given in the appendix.

Table 3 Exposures ( $\mu\text{g}/\text{m}^3$ -years benzo(a)pyrene) for which PC >50% (cases meeting exposure criterion (%))

Models of exposure and smoking	Linear model		Power curve model	
	Point estimate	Upper confidence limit	Point estimate	Upper confidence limit
Multiplicative model	342.2 (2.7)	190.1 (19.2)	210.3 (15.7)	45.9 (39.3)
Additive model:				
0 p-y	37.8 (40.2)	21.0 (46.7)	0.6 (78.1)	0.1 (83.7)
10 p-y	162.6 (22.2)	90.3 (32.5)	29.6 (42.3)	6.5 (61.5)
20 p-y	287.3 (6.2)	159.6 (22.8)	132.7 (26.6)	29.0 (42.9)
30 p-y	412.1 (1.8)	228.9 (13.0)	343.2 (2.7)	74.9 (34.0)
40 p-y	536.9 (0.9)	298.2 (4.7)	688.9 (0.3)	150.4 (23.4)
50 p-y	661.6 (0.3)	367.5 (2.7)	1194.8 (0.0)	260.8 (9.5)
60 p-y	786.4 (0.3)	436.8 (1.8)	1883.7 (0.0)	411.1 (1.8)
70 p-y	911.2 (0.0)	506.1 (1.5)	2776.6 (0.0)	606.1 (0.9)
80 p-y	1035.9 (0.0)	575.4 (0.9)	3893.9 (0.0)	849.1 (83.7)

The exposures are given by the formula  $(1/\beta)^{1/p}$  for the multiplicative model, and  $((1 + (0.33 \times \text{pack-years}))/9.05\beta)^{1/p}$  for the additive model. The exact parameter values used to obtain these figures are given in the appendix ( $p = 1$  for the linear model). p-y = pack-years.

50% PC exposure for the multiplicative model (342.2  $\mu\text{g}/\text{m}^3$ -years benzo(a)pyrene; 190.1 with the UCL) does not depend on smoking. Note that it is the same as that given by the formula for the additive model applied to a subject with average smoking history—24.4 pack-years. The qualifying exposures for the power curve model are lower than those for the linear model, for the multiplicative model, and for light smokers under the additive model, but higher for heavy smokers under the additive model. Table 3 also gives the proportion of the 338 cases in the study that exceeded the cumulative exposures shown, and thus would have been compensated if these had been used as minimum qualifying exposures.

For the additive model, the percentages show, for each level of smoking, the proportion of all 338 cases that exceed the exposure shown, regardless of their smoking habit. It was not possible to obtain actual smoking numbers, as the smoking information is unknown or of poor quality for many of them. However, as there was no evidence for association between smoking and exposure to coal tar pitch volatiles, the percentages should be good approximate estimates of the proportions after dividing subjects by exposure, if smoking habits are obtained without bias.

About the same overall proportion of cases would be compensated with a minimum exposure criteria based on the multiplicative model as one based on the additive model. Different people would be compensated, however, with the additive model favouring non-smokers or low smokers.

Although not listed on the tables, we also applied the approach of Chase *et al*<sup>13</sup> (see appendix). According to this, the "risk apportioned to coal tar pitch volatiles" is for non-smokers the same as the PC for the multiplicative model, but for smokers it reduces rapidly from this figure with pack-years. At just 10 pack-years, the risk apportioned to a case at the highest listed level of exposure (400  $\mu\text{g}/\text{m}^3$ -years) is 23%, and 36% if the upper 95% confidence limit is used. Thus, applying this criterion would result in many fewer cases being compensated than were caused by coal tar pitch volatiles, in the sense that they would not have occurred without exposure to coal tar pitch volatiles. A criterion lower than 50% could be used to remedy

this, in which case the approach would give results somewhat similar to those with the PC and the additive model for smoking and coal tar pitch volatiles, with slightly less contrast between smokers and non-smokers. We did not, in calculating pack-years, exclude those before 1966 (taken as the United States Surgeon-general's first clear warning of the hazards of smoking), as the significance of this date, at least in Canada, was unclear.

## Discussion

Allowing for all the uncertainties, we think that a causal association between exposure to coal tar pitch volatiles and lung cancer in the smelter studied is beyond reasonable doubt.<sup>2</sup> Given the similarity of the constituents of coal tar pitch volatiles in this and other Soderberg plants<sup>2</sup> it is highly likely that this is so for these plants too, although the level of risk will differ if the level of emissions is different. In the absence of other data, the exposure-response relations estimated from the Quebec smelter remain the best basis for estimating risk and assessing cause of disease after exposure in other Soderberg plants. Extrapolation of this relation to other workers exposed to PAHs is more speculative, but where other closer data do not exist, it may not be unreasonable, especially if the PAH profile is found to be similar.

We propose the PC as a basis for considering compensation claims, despite the critiques that have been mounted against it. The concept of apportioned risk advanced by Chase *et al*<sup>13</sup>, unlike the PC, is not a standard epidemiological one, and in particular does not correspond in any sense to the proportion of cases that would not have occurred without exposure. It seems to be an arbitrary alternative way of dividing the risk due to the combined action of smoking and asbestos so that smokers are less likely to obtain compensation, even if they have the same PC as a non-smoker. We can find no record of the approach being used in compensation for workers. Furthermore, its application here would compensate far fewer cases than were shown to be due to exposure to coal tar pitch volatiles in the epidemiological sense. Robins and Greenland's objections to the concept of the PC are, we have argued, only academic in the context that risks affect only a small proportion of exposed subjects. We therefore propose the PC as the best available practical aid to assessing cause of disease.

Table 4 Some possible compensation criteria

Basis	Criterion†	Cases compensated n (%)
1 The exposure that would have allowed compensation of the same proportion of cases studied that were attributed to coal tar pitch volatiles, according to: (a) a grouped analysis; (b) the linear model; and (c) the power curve model.	(a) $\geq 141.0^*$ (b) $\geq 212.3^*$ (c) $\geq 181.3^*$	86 (25.4) 50 (14.7) 69 (20.4)
2 Analysis by department: RR>2, PC>50%	$\geq 20$ years in Soderberg potroom or carbon plant	80 (23.7)
3 Analysis by cumulative exposure in groups RR>2, PC>50%	$\geq 100$	106 (31.4)
4 Linear model for benzo(a)pyrene: point estimate of $\beta$ ; RR>2, PC>50%	$\geq 342.2$ (-5 years*) (additive model depends on smoking)	9 (2.7)
5 Linear model for benzo(a)pyrene 95% upper confidence limit for $\beta$ ; RR>2, PC>50%	$\geq 190.1$ (-5 years) (additive model depends on smoking)	64 (19.2)
6 Power curve model for benzo(a)pyrene: point estimate of $\beta$ ; RR>2, PC>50%	$\geq 210.3$ (-5 years*) (additive model depends on smoking)	53 (15.7)
7 Power curve model for benzo(a)pyrene 95% upper confidence limit for $\beta$ ; RR>2, PC>50%	$\geq 45.9$ (-5 years*) (additive model depends on smoking)	133 (39.2)

\*Criteria 1 and 4–7 are based on mathematical models in which cumulative exposure occurring in the past five years before diagnosis was not counted. Criteria 2 and 3 are based on tabulation of risk, in which all exposure was counted.

†Criteria are measured in  $\mu\text{g}/\text{m}^3$ -years benzo(a)pyrene, except for criterion 2.

The estimate of PC is only as good as the exposure-response relation from which it is derived. Apart from sampling uncertainty, reflected in the confidence intervals, the exposure-response models are subject to uncertainty due to possible biases that afflict all epidemiological studies to some degree: uncontrolled confounding bias and information bias. We argued elsewhere<sup>2</sup> that our study was not immune from such bias, but it was unlikely to be so strong as to radically alter results.

The form of the model used for the exposure-response relation is usually influential in determining the PC. We are confident that the models used are compatible with the data, but are unable to discriminate between importantly different models empirically, and are unable to exclude other possible models. The model implicit in the tabulation of risk by exposure categories has arbitrary boundaries between categories, but has the advantage of simplicity. Of the mathematical models, we prefer the linear to the slightly better fitting power curve model, because it is simpler, and standard in research on carcinogens. Nevertheless, the power curve model serves as an illustration of the influence of choice of model in deciding PC. Between the additive and multiplicative models for the combined action of coal tar pitch volatiles and smoking, we prefer the multiplicative model because of its simplicity, but we cannot discriminate between the two by how well they fit the data.

If an all or nothing decision on compensation is required, we must choose, in addition to a model, a minimum exposure criterion. This choice is not solely a scientific one. We have assumed that there is particular interest in the exposure at which the PC is 50%, because of precedent in the Quebec and other workers' compensation boards. However, choice of another level of PC, compensation on a sliding scale proportional to the PC, or not using the PC, is no less "scientific".

Table 4 summarises several possible ways to identify cases, each scientifically defensible, for which the PC >50%. Unfortunately, different approaches lead to quite substantially different criteria. The approach summarised in the first row of table 4 takes as minimum exposure criterion that exposure which would have allowed compensation of the total proportion of cases studied that were attributed to coal tar pitch volatiles. The uncertainties due to model choice are relatively small. The lowest figure ( $141 \mu\text{g}/\text{m}^3$ -years benzo(a)pyrene) could perhaps be justified on the grounds of giving the benefit of doubt to the worker, and because of simplicity in its derivation. However, a weakness of this approach is its dependence on the particular workforce studied. The same criterion applied to another workforce, or even the same workforce in future years, would not compensate the "right" proportion of cases.

The second and third options, based on subjects grouped by exposure, are the simplest to explain. Such grouped analyses being used to decide compensation may be criticised as depending on choice of cut off points, which is arbitrary, and it often leads to many fewer cases being compensated than are attributable to work exposures. Here, however, the proportions compensated lie close to those estimated as attributable to coal tar pitch volatiles. Options 4–7 rely more on "black box" statistical techniques, but are arguably more rational in assuming plausible smooth exposure-response relations. The use of the additive models for smoking and coal tar pitch volatiles adds further complexity, and requires knowledge of smoking history of cases in pack-years, which may be difficult to ascertain accurately.

Given the uncertainties in all approaches, we suggest that if 50% PC is considered desirable, the simple approach listed as 3 in table 4 ( $>100 \mu\text{g}/\text{m}^3$ -years benzo(a)pyrene) is the most attractive. It can be justified simply, on the basis of published tables of risk.<sup>2</sup> It can be said to give a slight benefit of the doubt to workers in that it would be likely to compensate a somewhat higher fraction of cases than were actually caused by coal tar pitch volatiles. It is preferable to the also simple option 2 ( $> 20$  years in the potroom), if estimates of exposure to benzo(a)pyrene are available, because it would apply in other plants with different levels of exposure.

After considering the alternatives, the compensation board of the Canadian Province of Quebec propose to compensate all cancer cases for whom their PC is greater than 50%, according to the upper 95% confidence limit of a risk predicted by the linear RR model. The board prefer the additive model for smoking and effects of coal tar pitch volatiles, so that workers require less exposure to qualify for compensation if they smoked less. Proposed qualifying exposures are thus those given in the second column of table 3. For example, (in  $\mu\text{g}/\text{m}^3$ -years benzo(a)pyrene): never-smokers: 21.0; 20 pack-years: 159.6; 40 pack-years: 298.2; 60 pack-years: 436.8. Given typical past potroom exposures of about



10  $\mu\text{g}/\text{m}^3$  benzo(a)pyrene, this implies that non-smokers will be compensated even with a very short duration of average potroom exposure, moderate smokers if exposed for a long period, and heavy smokers probably not at all.

Others contributing to the work on which this report is based include D Baris, M Hodge, C Tremblay, and A Vandal. We were also influenced by an unpublished report by J Spinelli and P Band of the British Columbia Cancer Agency, and we are indebted to the workers and management of Alcan Aluminium for their cooperation.

#### Appendix: The combined effect of coal tar pitch volatiles (CTPV) and smoking

Our model for lung cancer relative risk (strictly hazard ratio) and exposure to CTPV, without considering smoking (expression 2 in the main text), was:

$$RR(x) = 1 + \beta x. \quad (A1)$$

To extend this, we denote cumulative exposure to CTPV as  $x_1$ , and pack-years of smoking as  $x_2$ . We consider two models: a linear multiplicative model:

$$RR(x_1, x_2) = (1 + \beta_1 x_1) (1 + \beta_2 x_2) \quad (A2)$$

and a linear additive model:

$$RR(x_1, x_2) = 1 + \gamma_1 x_1 + \gamma_2 x_2. \quad (A3)$$

There was evidence in our data that exposure to CTPV was not associated with smoking, but because of limitations in the quality of smoking information, we considered it undesirable to directly estimate the smoking model parameters from it.

For the multiplicative model, the absence of an association between smoking and exposure to CTPV allows us to use the estimate of  $\beta$  from model (A1) as an estimate of  $\beta_1$  in model (A2). An estimate of  $\beta_2$  is not required to make an estimate of PC, because the smoking effect cancels, giving:

$$PC(x) = \beta x_1 / (1 + \beta x_1), \quad (A4)$$

whatever the level of smoking.

For the additive model we require estimates of the parameters  $\gamma_1$  and  $\gamma_2$ . We used an estimate of  $\gamma_2$  from a large study of cancer carried out in Montreal.<sup>19</sup> This study was chosen from among several as the most likely to be relevant for our population, but results from other studies on smoking and lung cancer give similar results. The estimate of  $\gamma_2$  (increment in relative risk per pack-year) was 0.33, with 95% CI 0.20–0.62 (Siemiatycki, personal communication). We estimated  $\gamma_1$  from  $\beta$ , assuming that the average pack-years in study subjects is equal to 24.4, the mean among those with reasonable quality smoking data at the mean age at death (64 years). This implies a mean "baseline" lung cancer risk, given this smoking habit, of  $1 + (0.33 \times 24.4) = 9.05$  times that of a group comprised entirely of non-smokers. Thus the observed exposure-response relation for CTPV ( $RR(x) = 1 + \beta x$ ) may be taken to apply to a baseline risk 9.05 times that of non-smokers. The parameter  $\gamma_1$ , which can be thought of as the increment of relative risk per unit CTPV in a non-smoker, is thus 9.05 times the increment  $\beta$  (model A1) found for the study population as a whole:

$$\gamma_1 = 9.05\beta \quad (A5)$$

Risk of lung cancer in a subject with CTPV =  $x_1$ , and pack-years =  $x_2$ , relative to someone with the same smoking history but not exposed to CTPV ( $x_1 = 0, x_2$ ) is:

$$RR(x_1 | x_2) = RR(x_1, x_2) / RR(x_1 = 0, x_2) = (1 + \gamma_1 x_1 + \gamma_2 x_2) / (1 + \gamma_2 x_2) \quad (A6)$$

Thus:

$$PC(x_1 | x_2) = (RR(x_1 | x_2) - 1) / (RR(x_1 | x_2) - 1) = \gamma_1 x_1 / (1 + \gamma_1 x_1 + \gamma_2 x_2) \quad (A7)$$

Substituting  $\gamma_1 = 9.05\beta$  and  $\gamma_2 = 0.33$  yields expression (6) of the main text.

Analogous expressions for the power curve model are obtained by replacing  $\beta_1 x_1$  and  $\gamma_1 x_1$  by  $\beta'_1 x_1^p$  and  $\gamma'_1 x_1^p$ . The exact parameter estimates used to calculate the figures in the tables were:  $\beta = \beta_1 = 0.002922$  (0.001168–0.005261),  $\beta'_1 = 0.1313579$  (0.0595975–0.2340497),  $p = 0.3795$ ,  $\beta_2 = \gamma_2 = 0.33$ .

The calculation of "risk apportioned to CTPV" (RACTPV) from the linear multiplicative model, with the approach of Chase *et al.*,<sup>13</sup> is immediate from expression (2) and (A2):

$$RACTPV = \beta_1 x_1 ((1 + \beta_1 x_1)(1 + \beta_2 x_2) - 1) / ((1 + \beta_1 x_1)(1 + \beta_2 x_2)(\beta_1 x_1 + \beta_2 x_2)) \quad (A8)$$

- 1 Ronneberg A, Langmark F. Epidemiological evidence of cancer in aluminum reduction plant workers. *Am J Ind Med* 1992;22:573–90.
- 2 Armstrong B, Tremblay C, Baris D, Theriault G. Lung cancer mortality and polynuclear aromatic hydrocarbons: a case-cohort study of aluminium production workers in Arvida, Quebec, Canada. *Am J Epidemiol* 1994;139:250–62.
- 3 Armstrong BG, Tremblay C, Theriault G. Compensating bladder cancer victims employed in aluminum reduction plants. *J Occup Med* 1988;30:771–5.
- 4 Rothman KJ. Modern epidemiology. Boston: Little Brown, 1986.
- 5 Bond V. The cancer risk attributable to radiation exposure: some practical problems. *Health Phys* 1981;40:108–11.
- 6 Enterline PE. Sorting out multiple causal factors in individual cases. In: Chiaze L, Lundin F, ed. *Epidemiologic methods for occupational and environmental health studies*. Ann Arbor, MI: Ann Arbor Science Publications, 1983: 177–84.
- 7 Brennan TA, Carter F. Legal and scientific probability of causation of cancer and other environmental disease in individuals. *J Health Polit Policy Law* 1985;10:33–80.
- 8 Jenkins B. Memorandum opinion. In: Allen I, *et al.* *Plaintiff v United States of America, defendant*. Salt Lake City, Utah: US District Court for the District of Utah, Central Division, 1984:374–5.
- 9 Industrial Diseases Standards Panel. *Report to the workers' compensation board on the Ontario uranium mining industry*. Toronto: IDSP, 1989.
- 10 Industrial Diseases Standards Panel. *Report to the workers' compensation board on the Ontario gold mining industry*. Toronto: IDSP, 1987.
- 11 Lagakos SW, Mosteller F. Assigned shares in compensation for radiation-related cancers. *Risk Anal* 1986;6: 345–7.
- 12 Muir D. Compensating occupational diseases. *CMAJ* 1993;148:1903–5.
- 13 Chase GR, Kotin P, Crump K, Mitchell RS. Evaluation for compensation of asbestos-exposed individuals. *J Occup Med* 1985;27:189–98.
- 14 Robins JGS. Estimability and estimation of expected years of life lost due to a hazardous exposure. *Stat Med* 1991; 10:79–93.
- 15 Robins J, Greenland S. The probability of causation under a stochastic model for individual risk. *Biometrics* 1989;45: 1125–38.
- 16 Robins J, Greenland S. Estimability and estimation of excess and etiologic fractions. *Stat Med* 1989;8:845–59.
- 17 Greenland S, Robins JM. Conceptual problems in the definition and interpretation of attributable fraction. *Am J Epidemiol* 1988;128:1185–97.
- 18 Bjorseth A, Bjorseth O, Fjeldstad P. Polycyclic aromatic hydrocarbons in the work atmosphere. Determination of area-specific concentrations and job-specific exposure in a vertical pin Soderberg aluminium plant. *Scand J Work Environ Health* 1981;7:223–32.
- 19 Siemiatycki J, Dewar R, Krewski D, Desy M, Richardson L, Franco E. Are the apparent effects of cigarette smoking on lung and bladder cancers due to uncontrolled confounding by occupational exposures? *Epidemiology* 1994;5:557–65.